The concealed middle?

The concealed middle? An exploration of ordinary young people and school GCSE subject area attainment

Christopher J. Playford and Vernon Gayle

[a]*Administrative Data Research Centre, University of Edinburgh, Edinburgh*
[b] *School of Social and Political Science, University of Edinburgh, Edinburgh*

The General Certificate of Secondary Education (GCSE) is the standard qualification undertaken by pupils in England and Wales at the end of year 11 (age 15–16). GCSEs continue to play an important and central role in young people's educational and employment pathways. Within the sociology of youth, there has been recent interest in documenting the lives and educational experiences of 'ordinary' young people. In this paper, we analyse school GCSE attainment at the subject area level. This is an innovative approach and our motivation is to explore substantively interesting patterns of attainment that might be concealed in analyses of overall attainment or attainment within individual subjects. We analyse data from the Youth Cohort Study of England and Wales using a latent variable approach. The modelling process uncovered four distinctive latent educational groups. One latent group is characterised by high levels of overall attainment, whereas another latent group is characterised by poor GCSE performance. There are two latent groups with moderate or 'middle' levels of GCSE attainment. These two latent groups have similar levels of agglomerate attainment, but one group performs better in science and the other performs better in arts GCSEs.

Keywords: educational attainment; GCSE; missing middle; sociology of youth; youth cohort study of England and Wales; latent class models

1. Introduction

Historically, school examination results were a private matter and the awareness of results day was usually confined to pupils, teachers and parents. School exam results are now an annual newsworthy item in Britain (for example, see Joint Council for Qualifications 2012; Chalabi 2013; Lim 2013). Every summer the British media transmit live broadcasts of groups of young people receiving their grades (Warmington and Murphy 2004). This recurrent event illustrates, and reinforces, the importance of school-level qualifications in Britain.

The General Certificate of Secondary Education (GCSE) is the standard qualification undertaken by pupils in England and Wales at the end of year 11 (age 15–16; Department for Education 1985; Mobley et al. 1986; North 1987). We consider that school GCSE attainment is worthy of sociological attention because these qualifications mark the first major branching point in a young person's educational career and play a critical role in determining post-compulsory education and employment pathways. School GCSEs are often the only qualifications achieved by pupils who leave education at the minimum age

(Leckie and Goldstein 2009). School GCSE attainment is strongly related to participation in post-compulsory education (Payne 1995, 2000, 2001, 2003). Rice (1999) reports a clear relationship between poor school GCSE performance, unemployment and low levels of participation in further education. Babb (2005) concludes that young people's experiences at school and their attainment at GCSE level are strong determinants of their future success in both education and employment. Through the detailed examination of household panel data, Murray (2011) similarly observes that the negative effects of poor GCSE attainment follow young people into early adulthood. More generally, Jones et al. (2003) illustrate that workers with poor school-level qualifications (e.g. GCSEs) generally have less favourable long-term outcomes in the labour market.

University entry in Britain is primarily based on attainment in advanced level qualifications gained at age 18+, for example, General Certificate of Education Advanced Levels. In addition, universities may ask for a specific number of GCSEs, and some courses require higher grade GCSEs. GCSE English at grade C or above is a universal requirement for university entry in Britain. For many courses a grade C GCSE or above in Mathematics is often required. Some British university courses require higher grade passes in GCSE English and Mathematics, and a grade C or above in a foreign language. Some courses require that the pupil achieves certain GCSE grades and subjects in a single sitting and do not accept resits (The Russell Group 2013).

There is a growing interest in examining attainment in individual GCSE subjects. Jin, Muriel, and Sibieta (2011) assert that the subjects and courses young people choose to take from age 14 onwards can have profound implications for future education and economic opportunities. There is particular interest in GCSE attainment in English and Mathematics, and these subjects have always been regarded as a core part of the educational curriculum. A key recommendation of the Wolf Report (2011) is that pupils who have not secured a good pass in English and Mathematics GCSEs should continue to study these subjects.[1] This position has received both positive and negative attention in the media (see Garner 2013). There has also been wider interest in GCSE subject choices in schools (see Clark 1995; Adey and Biddulph 2001; Murphy and Whitelegg 2006; Lamont and Maton 2008; Sullivan, Zimdars, and Heath 2010; Pau, Hall, and Grace 2011; Weedon 2011; Institute of Physics 2012, 2013; Owen-Jackson 2013; Mujtaba and Reiss 2013).

In this paper, we turn our attention to exploring school GCSE attainment at the subject area level, rather than at the level of the individual GCSE subject. This is an innovative approach to studying school GCSE attainment. Our motivation is that we speculate that there might be substantively interesting patterns of attainment comprising individual GCSE subjects. These patterns will be masked when the focus is either overall attainment or attainment in individual subjects.

2. Theorising ordinary young people

A further motivation for studying school GCSE attainment at the subject area level is to begin to better theorise the experiences of ordinary pupils who perform somewhere between the obviously successful and unsuccessful levels of achievement. The notion of 'ordinary kids' is far from novel and is employed in youth research (for example, see Jenkins 1983; Coles 1986; Brown 1987; Pye 1988; France 2007). The idea of returning analytical attention to the lives of ordinary youths is currently in vogue, and a recent special issue of *Sociological Research Online*[2] was devoted to studying ordinary young

people. Roberts (2012) comments that 'ordinariness' has largely epitomised a 'missing middle' in youth research, particularly when examining young people's experiences of education.

Roberts (2011, 22) appeals to youth researchers to better document the experiences of ordinary young people through the secondary analysis of large-scale data-sets to establish their social characteristics and how well qualified they might be. This appeal to explore 'ordinary' youth using social survey data has recently been taken up. Connelly, Murray, and Gayle (2013) explored the middle ground between what can broadly be termed as the educationally successful and the unsuccessful, using data from the British Household Panel Survey. They identified a group of pupils with what might plausibly be termed 'middle' levels of school GCSE attainment, and they examined this group's activities in early adulthood. Through the application of statistical models, they compared this group's education and economic activities with the activities of their more and less educationally successful peers. The 'middle' group differed in their economic activities in early adulthood and notably made the transition from education into employment earlier. Connelly, Murray, and Gayle (2013) also undertook more exhaustive analyses and concluded that there was not a clearly defined 'middle' group, and they warned against making extended claims about this group without further exploration.

Following on from this research, in another paper Gayle, Murray, and Connelly (2014) undertook a replication analysis using the Youth Cohort Study of England and Wales (YCS), with the intention of augmenting, and therefore extending, their earlier work with detailed secondary analysis of school attainment using specialist youth data. The overall message was that educational inequality persisted in school GCSE attainment throughout the 1990s and into the early 2000s. They reported that the overall pattern of school examinations has been one of increasing performance, but GCSE attainment remained highly stratified. In general girls performed better than boys, and there were some marked differences in attainment for pupils from the main minority ethnic groups. A striking result was the impact of parental socioeconomic positions, and to a lesser extent other variables measuring the young person's home environment.

Gayle, Murray, and Connelly (2014) state that they began their analysis with an open mind, but their explorations failed to convince them that there were distinctive, or discrete, categories of GCSE attainment. The evidence failed to persuade them that there were crisp boundaries that demark a 'middle' category of moderate GCSE attainment. The analyses persuaded the authors that GCSE attainment was situated on a continuum. Whether measured by the number of GCSEs gained at higher grades, or by computing a GCSE score, attainment was similarly stratified. With the exception of the sharp spike of young people that were unsuccessful in gaining any higher grade GCSEs, they failed to observe the presence of any clusters that indicated clear cohesive GCSE attainment groups.

Merton (1987) cautioned that before sociologists proceed to explain or to interpret a phenomenon, it is advisable to establish that the phenomenon actually exists, and that it is enough of a regularity to require and to allow explanation. Mindful of this methodological prescription the overall goal of this paper is to explore patterns of school GCSE attainment at the subject area level with the aim of establishing if there is evidence that there are groups of young people with 'middle' levels of attainment. To borrow a term from outside of sociology, the overall motivation of this paper is therefore to establish 'proof of concept'.

3. Data

The Youth Cohort Study of England and Wales (YCS) is a major longitudinal study that began in the mid-1980s. It is a large-scale nationally representative survey funded by the government and is designed to monitor the behaviour of young people as they reach the minimum school leaving age and either remain in education or enter the labour market. The YCS has been successfully used to explore the relationship between parental and family backgrounds and filial attainment (Drew, Gray, and Sime 1992; Drew 1995; Demack, Drew, and Grimsley 2000; Gayle, Berridge, and Davies 2003; Connolly 2006; Gayle, Lambert, and Murray 2009a; Sullivan, Heath, and Rothon 2011).

The YCS survey collects detailed information on the young person's qualifications and experiences of education, as well as information on employment and training. A limited amount of information is collected on the young person's personal characteristics, and their family and circumstances at home. The YCS sample is nationally representative of Year 11 pupils in England and Wales. A large sample from an academic year group (a cohort) is contacted in the spring following Year 11. The young people are usually of age 16–17 when they are first contacted. The main data collection instrument is a postal questionnaire. The cohort members are usually re-contacted and surveyed on at least two further occasions (for example at ages 17–18 and 18–19).

The Youth Cohort Study is primarily a monitoring tool and was not specifically designed for social science research. There are a number of challenges associated with analysing YCS data, most notably inadequate documentation of the procedures used to construct the data-sets (Croxford 2006). Therefore to start this investigation we focused on a single cohort of pupils. We restricted our analyses to YCS Cohort 6 who were surveyed in 1992 (Youth Cohort Study of England and Wales, 1992–1994). YCS Cohort 6 is an appropriate choice because it is a bigger YCS cohort (n = 14,576 pupils). YCS Cohort 6 is a suitable cohort to begin exploratory analyses upon because GCSEs had been undertaken for a number of years and teething problems associated with teaching and examinations should have largely been expunged. We confine the analyses to young people who attended comprehensive schools in Year 11. These pupils were at non-fee paying State funded schools that were non-selective. We undertake single-level analyses because there are no school-level or Local Authority-level indicators deposited with the data-set.

Measuring school GCSE attainment

The question of how to measure education and qualifications, or indeed what 'measure' means, raises interesting issues since there is no agreed standard way of categorising educational qualifications (Prandy, Unt, and Lambert 2004; Schneider 2011). GCSEs are the standard school qualification undertaken by pupils in England and Wales; they are usually a mixture of assessed coursework and examinations (Ashford, Gray, and Tranmer 1993). GCSE subjects are assessed separately and a subject-specific GCSE is awarded. It is usual for pupils in Years 10 and 11 to study for about nine subjects, which will include core subjects (e.g. English, Mathematics and Science) and non-core (optional) subjects.[3,4]

Each GCSE subject is awarded a grade, historically the highest being grade A and the lowest grade G. From 1994 a higher grade of A* was introduced (Yang and Woodhouse 2001), but this new grade postdates the pupils in YCS Cohort 6. Because GCSEs are taken as diet of many subjects and each subject is awarded an alphabetical grade there is no obvious single, or agreed, measure of overall school GCSE attainment.

The attainment of five or more GCSEs at grades A–C, and now at grades A*–C, is a standard benchmark, for example in school performance league tables (Leckie and Goldstein 2009). The A*–C measure is routinely employed in a wide variety of social science applications (e.g. Gayle, Berridge, and Davies 2003; Babb 2005; Connolly 2006; Sullivan, Heath, and Rothon 2011; Tunstall 2011). The overall limitation of the measure is that it treats an A* in music, a B in Maths and a C in sociology equally in determining whether or not a pupil has five GCSEs at grades A*–C (Gorard and Taylor 2002). For quite some time, the government league tables have also included a measure of the proportion of pupils gaining five or more GCSEs at grades A*–C including Mathematics and English (Taylor 2011). The addition of achieving grades A*–C in Mathematics and English does not however overcome the more general obstacle of how best to suitably combine GCSE subjects results.[5]

In this analysis we constructed a range of measures of a pupil's overall school GCSE attainment. These measures include the benchmark measure of 5+ grades at A–C, and the number of passes at grades A–C. A plausible course of action is to construct a measure based on scores. There are an infinite number of possible scores that could be assigned to the alphabetical grades ascribed to the levels of GCSE attainment. Following Croxford, Ianelli, and Shapira (2007, 52), we calculated a measure of GCSE attainment based on allocating 7 points for an A*/A, 6 points for a B, 5 points for a C, 4 points for a D, 3 points for an E, 2 points for a F and 1 point for a G. This was in line with the earlier Qualifications and Curriculum Authority (QCA) scoring method.[6]

We constructed a capped measure of GCSE attainment to limit the effects of pupils achieving higher scores simply as a function of having taken more GCSEs. Webber and Butler (2007) use a similar approach on the advice of DfES officials. Our measure is capped at 84 points the equivalent of 12 GCSEs at grade A*/A. More recently, some official statistics are capped at the level of the best eight GCSEs, although other alternative approaches could be employed.

Haque and Bell (2001) convert GCSE attainment into numerical scores (A* = 8, A = 7 … U = 0) and calculate a mean GCSE score for each pupil. They chose this approach because they believe that this helps to prevent discrimination against pupils who have taken fewer GCSEs as a result of school policy. Similarly, we can envisage the use of other summary measures of attainment, for example median scores. The QCA now work with a different scoring system which awards an A* 58 points, an A 52 points, a B 46 points, a C 40 points, a D 34 points, an E 28 points, a F 22 points and a G 16 points.[7] We suspect that because the new and old scores for each GCSE grade are similarly spaced, the overall substantive interpretations of analyses that use the new scoring system will not be dramatically altered.

The resounding message is that there is no single clearly recognised, or agreed upon, overall measure of GCSE attainment. We use the term 'agglomerate' measures to describe these overall or summary measures of school GCSE attainment. Table 1 presents a series of agglomerate GCSE attainment measures for YCS Cohort 6.

5. Latent variable analyses

Pupils in England and Wales can choose options from a wide range of GCSE subjects. These choices include many modern languages, humanities, social sciences and arts related subjects. Given that each pupil studies for a mixture of core and non-core subjects, there are an extremely large number of subject combinations within the diet of school

Table 1. Agglomerate measures of GCSE attainment: YCS Cohort 6.

Summary measures	
% 0 A–C Passes	16
% 1–4 A–C Passes	36
% 1–4 A–C Passes (including English and Maths)	5
% 5+ A–C Passes	48
% 5+ A–C Passes (including English and Maths)	41
Mean GCSE points score (A)	38.0
Mean number of GCSEs studied (B)	8.1
Mean GCSE points score per GCSE studied (A/B)	4.7
Grade of mean points score per GCSE studied	C/D
Mean number of A–C passes	4.5
Mean number of A–F passes	8.1

Note: n = 14,281; weighted data.

GCSEs. In addition, some GCSEs (e.g. history and geography) are extremely common choices whereas other GCSEs (e.g. Latin) are less popular (for example, see Gill 2012, Table 11). Therefore, adopting an approach that reasonably reduces the complexity of GCSE combinations is necessary in order to illuminate patterns of GCSE performance.

The GCSE subjects undertaken by pupils in YCS Cohort 6 have been simplified into five main groups from the 17 most popular GCSEs (see Table 2). Given the large number of subjects and potential subject combinations, a process of simplification is essential to operationalise the analysis. The seven category alphabetical grading scheme (A–G) also requires reorganisation into a numerical scheme. We code individual GCSEs into higher grades A–C and lower grades D–G. This categorisation was chosen because it is routinely used within the education system, official statistics, educational research and by employers. The observed patterns of attainment further convince us that this is a sensible approach.[8]

We use a latent classification approach for the analysis of subject area school GCSE attainment. Latent class models relate a set of observed (usually discrete) variables to a set of latent or unmeasured classes (McCutcheon 1987, 1996; Becker and Yang 1998;

Table 2. Year 11 school GCSE subject areas, YCS Cohort 6.

Subject groups	GCSEs	Numberof pupils gaining A–Caward
English	English	10,352
Maths	Mathematics	8284
Science	Biology; Physics; Chemistry; Double Science; Other Science	7807
Humanity	History; Geography; Other humanity; Religious education	9151
Other subject	French; CDT; Other language; Arts; Physical education; Other GCSEs	10,130
Total		14,281

Note: The 17 most frequently undertaken GCSEs; unweighted data.

Table 3. Year 11 School GCSE attainment: tetrachoric correlations.

	English	Mathematics	Science	Humanity	Other subject
English	1.00				
Mathematics	0.98	1.00			
Science	0.96	0.95	1.00		
Humanity	0.99	0.96	0.94	1.00	
Other subject	0.99	0.95	0.94	0.97	1.00

Note: n = 14,281; unweighted data, YCS Cohort 6.

McCutcheon 2002). Latent class models are especially suitable for analyses of a set of categorical observed measures that are highly interrelated (McCutcheon 2002, 56). A major attraction of this modelling approach is that it allocates individuals to latent classes based on the observed variables (Bartholomew et al. 2008, 272). These models have an obvious appeal for the study of multiple GCSEs which are measured on a categorical scale, and which may be strongly correlated.

After estimating an appropriate set of latent class models, we investigate the observed characteristics associated with membership of the latent class. The analyses are progressed in the spirit of exploratory data analysis (see Tukey 1977; Marsh and Elliott 2008). Because the term 'class' is used within education to denote forms or classrooms, and is also used as a short-hand for social class, from this point onwards we use the term 'latent group' rather than 'latent class' to avoid confusion.

Performance in individual GCSEs is highly correlated. The tetrachoric correlations are reported in Table 3.[9] For example, gaining an A–C pass in Mathematics is highly correlated with gaining an A–C pass in Science (0.95). The correlation between achieving an A–C grade in Science and an A–C in Humanities is slightly weaker (0.94).

6. School GCSE attainment and latent educational groups

Table 4 reports the summary statistics for a series of latent group models.[10] The four group model is an improvement on the two and three group models, with a lower deviance and associated degrees of freedom.[11] The four group model is preferred to the five group model because it is the most parsimonious model, with the lowest AIC and BIC measures.

The results of the four category latent group model are reported in Table 5. In this model, there are five observed variables (GCSE subject groupings which are labelled English, Maths, Science, Humanity and Other). Each variable has two outcomes, an A–C pass or a D–G pass. A single latent categorical variable is estimated. The label ascribed to each of the latent groups describes the overall substantive pattern of GCSE attainment.

The model identifies four latent educational groups and assigns pupils. The posterior (group) and prior (item) probabilities estimated in the modelling process are reported as percentages in Table 5 (for a full discussion see Bartholomew et al. 2008, 273). About 51% of pupils are assigned to group 1, 8% to group 2, 18% to group 3 and 23% to group 4. We ascribe the labels 'Good Grades', 'Science', 'Arts' and 'Poor Grades' based on the overall school GCSE attainment profiles of the pupils in these latent groups. Latent group 1 we label as 'Good Grades' because this group is characterised by good attainment

Table 4. Latent educational group models (goodness-of-fit statistics).

Model	χ^2	Deviance	df	ΔDeviance	AIC	BIC
Group 2	1208.9	1086.7	20	–	1046.7	895.4
Group 3	231.2	199.2	14	888	171.2	65.3
Group 4	5.0	5.0	8	194	−10.9	−71.5
Group 5	0.3	0.3	2	5	−3.7	−18.8

Note: n = 14,281, YCS Cohort 6.

across all of the subject areas. We label latent group 4 as 'Poor Grades' because this group is characterised by poor GCSE attainment across all of the subject areas.

There are two latent groups with distinctive educational profiles that locate between good and poor performance. We label the first of these, latent group 2, as 'Science'. This is because pupils within this latent group have good levels of attainment in science, and 75% pass a science GCSE with A–C grades. Latent group 3 we label as 'Arts'. Pupils within this latent group have high levels of attainment in GCSE English, with 83% passing with A–C grades. Notably, pupils in latent group 3 perform relatively poorly in GCSE Maths and only 32% pass with A–C grades. They also perform poorly in science and only 14% achieve a science GCSE pass with A–C grades.

Table 6 reports a series of agglomerate measures of GCSE attainment that are observed for each latent group. Pupils in latent group 1 (Good Grades) perform well on all agglomerate measures of school GCSE attainment. By contrast, pupils in latent group 2 and latent group 3 perform less well than their counterparts in latent group 1 but better than their counterparts in latent group 4. Latent group 2 (Science) and latent group 3 (Arts) can reasonably be considered as groups with moderate or 'middle' levels of school GCSE attainment. An important early finding is that pupils in the 'Science' and the 'Arts'

Table 5. Latent group model results (four group model) school GCSE attainment – posterior probabilities and prior probabilities (percentages), YCS Cohort 6.

Latent group	1 Good grades	2 Science	3 Arts	4 Poor grades
Posterior probabilities				
Percentage assigned to group (%)	51	8	18	23
Prior probabilities				
English A–C (%)	98	49	83	17
English D–G (%)	2	51	17	83
Maths A–C (%)	94	56	32	6
Maths D–G (%)	6	44	68	94
Science A–C (%)	91	75	14	5
Science D–G (%)	9	25	86	95
Humanity A–C (%)	95	47	59	9
Humanity D–G (%)	5	53	41	91
Other A–C (%)	93	62	71	27
Other D–G (%)	7	38	30	73
N	7268	1207	2573	3233

Note: All pupils gaining a GCSE pass at grades A–G; n = 14,281; Probabilities reported as percentages.

Table 6. Agglomerate measures of school GCSE attainment by latent group.

Latent group	1 Good grades	2 Science	3 Arts	4 Poor grades	All
Mean number of A–C passes	7.7	4.0	3.3	0.5	4.5
Mean number of A–F passes	8.8	8.0	8.2	6.9	8.1
Mean GCSE points score (A)	49.7	35.5	34.6	23.4	38.0
Mean number of GCSEs studied (B)	8.5	8.0	8.2	7.6	8.1
Mean points score per GCSE studied (A/B)	5.8	4.4	4.2	3.0	4.7
Grade of mean points score per GCSE studied	B	D	D	E	C/D

Note: All pupils gaining a GCSE pass at grades A–G, *n* = 14,281, weighted data, YCS Cohort 6.

groups have similar levels of agglomerate GCSE attainment, despite having different patterns of attainment at the subject area level.

Table 7 reports additional detailed GCSE attainment information for each latent educational group. The majority of the pupils in latent group 1 (Good Grades) achieve the longstanding benchmark of 5+ GCSEs at grades A–C. A similarly high percentage of pupils in this latent group achieve the more stringent standard of 5+ GCSEs at grades A–C including Maths and English. Consistent with their overall profile of performance, few pupils in latent group 4 (Poor Grades) achieve 5+ GCSEs. It is notable however that 13% of pupils in latent group 4 pass GCSE English at grades A–C, but only 4% pass Mathematics at grades A–C.

The educational performance of the two 'middle' groups is especially interesting. About 31% of pupils in latent group 2 (Science) achieve 5+ GCSEs at grades A–C. However, 44% of this group pass Maths at grades A–C, but only 39% pass English GCSE at this level. About 15% of pupils in latent group 3 (Arts) achieve 5+ GCSEs at grades A–C. But in contrast to pupils in latent group 2 (Science), 81% of pupils in latent group 3 (Arts) pass GCSE English at grades A–C. The 'Arts' latent group perform poorly in GCSE Mathematics, however, and only 33% pass with grades A–C.

There are clearly two latent groups of pupils with moderate or 'middle' levels of school GCSE attainment. The 'Science' latent group is much smaller than the 'Arts'

Table 7. Benchmark school GCSE attainment by latent group (column percentages).

Latent group	1 Good grades (%)	2 Science (%)	3 Arts (%)	4 Poor grades (%)	All (%)
GCSE English (Grade A–C)	100	39	81	13	66
GCSE Maths (Grades A–C)	93	44	33	4	51
Both GCSE English and Maths (Grades A–C)	93	8	21	0	45
5+ A–C Passes (any subject)	97	31	15	<1	47
5+ A–C Passes (including English and Maths)	91	2	5	0	40

Note: All pupils gaining a GCSE pass at grades A–G, *n* = 14,281, weighted data.

latent group. Pupils in these two 'middle' latent groups have similar levels of agglomerate GCSE attainment but their patterns of subject area attainment are substantially different. These underlying patterns are hidden without the application of the latent variable analyses.

7. Investigating the characteristics of the latent educational groups

In the next stage of the analysis, we investigate the observed characteristics associated with membership of the latent educational groups. Table 8 reports the characteristics of

Table 8. Characteristics of the latent educational groups (row percentages).

Latent group	1 Good grades	2 Science	3 Arts	4 Poor grades		*n*
Parental occupation NS-SEC						
1.2. Higher professional	68%	8%	12%	13%	100%	1279
1.1. Large employers and higher managerial	56%	6%	17%	21%	100%	1109
2. Lower managerial and professional	53%	8%	17%	22%	100%	4297
3. Intermediate	41%	9%	21%	29%	100%	2565
4. Small employers and own account	31%	11%	24%	34%	100%	1106
5. Lower supervisory and technical	33%	8%	21%	38%	100%	1258
6. Semi-routine	26%	9%	21%	43%	100%	1684
7. Routine	24%	9%	19%	48%	100%	983
Gender						
Girls	47%	5%	23%	25%	100%	7794
Boys	40%	12%	15%	34%	100%	6487
Housing						
Owned	47%	8%	19%	27%	100%	12,527
Rented	22%	10%	21%	48%	100%	1589
Other housing	45%	9%	18%	28%	100%	165
Household type						
Lives with both parents	44%	9%	18%	29%	100%	12,668
Only lives with mother	37%	8%	23%	32%	100%	1304
Only lives with father	30%	11%	19%	39%	100%	302
Other household	55%	10%	0%	36%	100%	7
Ethnicity						
White	43%	9%	19%	29%	100%	13,413
Black	34%	2%	31%	31%	100%	226
Indian	43%	12%	19%	27%	100%	323
Pakistani	38%	6%	26%	31%	100%	111
Bangladeshi	23%	9%	30%	39%	100%	38
Other Asian	57%	12%	17%	15%	100%	126
Other ethnicity	45%	14%	23%	19%	100%	44
Observations	44%	9%	19%	29%		100%
	6205	1220	2696	4160		14,281

Note: All pupils gaining a GCSE pass at grades A–G; *n* = 14,281; weighted data; YCS Cohort 6.

each of the latent groups. We estimate a multinomial logistic regression model which includes measures of parental occupation (measured by the National Statistics Socio-economic Classification NS-SEC), gender, housing tenure, household type and ethnicity as explanatory variables. These variables have been routinely used in previous studies of GCSE attainment (for example, Drew, Gray, and Sime 1992; Drew 1995; Demack, Drew, and Grimsley 2000; Gillborn and Mirza 2000; Gayle, Berridge, and Davies 2003; Connolly 2006; Gayle, Lambert, and Murray 2009b; Phillips 2009; Sullivan, Heath, and Rothon 2011; Connelly, Murray, and Gayle 2013). The results of the multinomial logistic regression model are reported in Table 9. The outcome variable is the latent educational group that the pupil has been assigned to.

Parents' socioeconomic position, measured by the National Statistics Socio-economic Classification (NS-SEC), has a strong overall effect ($p < 0.01$). This result is consistent with earlier analyses of GCSE attainment using the YCS (Drew, Gray, and Sime 1992; Drew 1995; Demack, Drew, and Grimsley 2000; Gayle, Berridge, and Davies 2003; Connolly 2006; Gayle, Lambert, and Murray 2009a; Playford 2011; Sullivan, Heath, and Rothon 2011). Parental socioeconomic positions are the most important predictor of latent educational group membership. Pupils with parents in less advantaged socioeconomic groups are generally more likely to be in a latent group other than latent group 1 (Good Grades). Most notable are the increased log odds of pupils with parents in lower supervisory and technical occupations (NS-SEC 5), semi-routine occupations (NS-SEC 6) and routine occupations (NS-SEC 7) being assigned to latent group 4 (Poor Grades).

Pupils with parents in intermediate occupations (NS-SEC 3), small employers and own account workers (NS-SEC 4), lower supervisory and technical occupations (NS-SEC 5), semi-routine occupations (NS-SEC 6) and routine occupations (NS-SEC 7) all have increased log odds of being in either latent group 2 (Science) or latent group 3 (Arts). It is notable that there are no significant differences in the parental socioeconomic effect for pupils in latent group 2 (Science) and latent group 3 (Arts) (see Figure 1). The subtlety of this effect could not have been detected without the classification of the latent educational groups, and this comparison could not have been readily illustrated without the calculation of quasi-variance based comparison intervals (see Gayle and Lambert 2007).

On reflection we consider that having data with a detailed measure of parental socioeconomic positions is important as it facilitates more comprehensive analyses. The NS-SEC measure has a reasonably high degree of resolution compared with more crude measures such as eligibility for free school meals. Eligibility for free school meals is often used as a proxy measure for the socioeconomic position of a pupil's family, and often it is the only suitable proxy available in administrative data-sets. Authors such as Rollock et al. (2015) indicate the limitations of free school meals as a measure. Eligibility for free school meals may work well in some research applications; however, the results reinforce our conviction that using more detailed socioeconomic measures will always be more desirable in educational research.

The underachievement of boys in GCSEs is well documented. For a review, see Department for Education and Skills (2007) and for more recent evidence, see House of Commons Education Committee (2014) and Department for Education 2015). It is worth noting that contemporary administrative data reveal that there is little gender difference in the uptake of some GCSEs, for example 32% of both male and female pupils undertake history GCSE. By contrast, home economics (child development) is undertaken by 7% of girls but less than 1% of boys. Conversely, physical education/sports studies is undertaken by 23% of boys, but only 12% of girls (see Gill 2012).

Table 9. Multinomial logistic regression model results – latent educational group membership.

	Good grades	Science		Arts		Poor grades	
	Coefficient	Coefficient	SE	Coefficient	SE	Coefficient	SE
Parental occupation NS-SEC							
1.2. Higher professional	0.00	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
1.1. Large employers and higher managerial	0.00	−0.01	(0.20)	0.58***	(0.15)	0.67***	(0.15)
2. Lower managerial and professional	0.00	0.33*	(0.15)	0.62***	(0.12)	0.75***	(0.13)
3. Intermediate	0.00	0.69***	(0.16)	1.05***	(0.13)	1.24***	(0.13)
4. Small employers and own account	0.00	1.11***	(0.18)	1.43***	(0.15)	1.68***	(0.15)
5. Lower supervisory and technical	0.00	0.79***	(0.18)	1.26***	(0.14)	1.72***	(0.14)
6. Semi-routine	0.00	1.13***	(0.17)	1.44***	(0.14)	2.02***	(0.14)
7. Routine	0.00	1.05***	(0.20)	1.38***	(0.16)	2.14***	(0.15)
Gender							
Girls	0.00	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Boys	0.00	1.07***	(0.08)	−0.21***	(0.06)	0.56***	(0.05)
Housing							
Owned	0.00	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Rented	0.00	0.80***	(0.12)	0.56***	(0.10)	0.98***	(0.09)
Other housing	0.00	0.14	(0.34)	0.13	(0.28)	0.15***	(0.23)
Household type							
Lives with both parents	0.00	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Only lives with mother	0.00	−0.10	(0.14)	0.25***	(0.10)	−0.01	(0.10)
Only lives with father	0.00	0.55**	(0.24)	0.35*	(0.20)	0.53***	(0.18)
Other household	0.00	−0.27	(1.11)	−12.79***	(0.52)	−0.19	(1.07)
Ethnicity							
White	0.00	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
Black	0.00	−1.05***	(0.36)	0.60***	(0.22)	0.19	(0.24)
Indian	0.00	0.30	(0.25)	−0.01	(0.19)	−0.18	(0.19)
Pakistani	0.00	−0.37	(0.45)	0.46	(0.29)	0.07	(0.31)
Bangladeshi	0.00	0.25	(0.73)	1.06*	(0.62)	0.39	(0.53)
Other Asian	0.00	−0.02	(0.41)	−0.35	(0.31)	−1.07***	(0.31)
Other ethnicity	0.00	0.33	(0.54)	0.19	(0.46)	−0.53	(0.66)
Constant	0.00	−2.88***	(0.14)	−1.73***	(0.11)	−2.00***	(0.12)
Observations	14,281						
AIC	2.05			Log Likelihood		−14,577.7	
BIC	−1172.86			Pseudo *R* Squared		0.052	

Note: Log Likelihood is estimated for unweighted data. Pseudo *R* Squared reported for weighted data. Survey weighted data with linearised standard errors; *$p < .10$; **$p < .05$; ***$p < .01$.

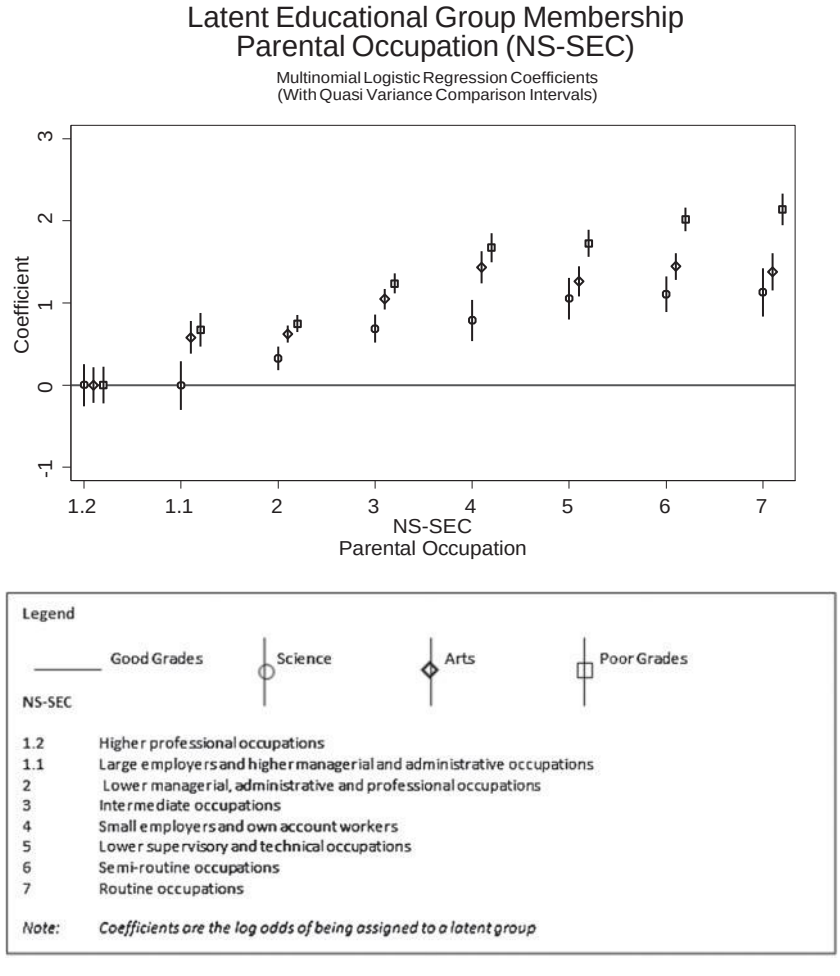

Figure 1. Multinomial logistic regression model of latent group member: Parental NS-SEC Coefficients.
Note: Other variables included in the model: gender, housing, household type, ethnicity, (see Table 9); weighted data; n = 14,576, YCS Cohort 6.

There is a significant overall gender effect ($p < 0.01$); however, a more nuanced effect is uncovered from the multinomial logistic regression model. What is most striking in the present analysis is that being male has a positive effect on membership of the 'Science' latent educational group but a negative effect on membership of the 'Arts' latent group. To investigate this further, and following Connelly, Murray, and Gayle (2013), we estimated a multinomial regression model of agglomerate school GCSE attainment, with three categories (5+ GCSEs at grades A–C, 1–4 GCSEs at grades A–C and no GCSE passes at grades A–C).[12] In the model of agglomerate GCSE attainment, there is an overall gender effect. Male pupils have higher log odds of achieving either 1–4 passes at grades A–C, or no GCSE passes at grades A–C, rather than of attaining 5+ GCSEs at grades A–C. Modelling the membership characteristics of the two 'middle' latent educational groups has exposed a more nuanced, and empirically informative, gender effect.

Overall the effects of ethnicity are significant but weak ($p < 0.01$). The ethnicity effect follows the usual pattern that some minority ethnic groups perform better than white pupils whilst some others perform less well. The model reports a significant positive effect for pupils of Indian origin being assigned to latent group 2 (Science). Most notably pupils in the Black minority ethnic group have lower log odds of being in latent group 2 (Science) but higher log odds of being in latent group 3 (Arts). By contrast, when we estimated a multinomial regression model of agglomerate school GCSE attainment, with three categories (5+ GCSEs at grades A–C, 1–4 GCSEs at A–C grades and no GCSE passes at grades A–C), the Black ethnicity effect is insignificant.[13] Modelling the membership characteristics of the two 'middle' latent educational groups has exposed a more nuanced ethnicity effect for Black pupils.

The familiar, but sociologically dispiriting result that despite decades of comprehensive education overall attainment is stratified, is evident for this cohort. Membership of the latent educational groups is also highly stratified. More socially advantaged pupils are more likely to be assigned to group 1 'Good Grades'. In contrast, less socially advantaged pupils are more likely to be assigned to group 4 'Poor Grades'. The analyses uncovered two latent educational groups with similar levels of moderate agglomerate school GCSE attainment but different overall patterns of subject level attainment. A notable new finding is that latent educational group 2 'Science' had a different gender profile to group 3 'Arts', but the parental socioeconomic classification profile for these two latent educational groups was similar.

8. Conclusion

The analyses presented above answer the appeal made by Roberts (2011, 22) for researchers to better document the experiences of ordinary young people through the secondary analysis of large-scale data-sets. This paper is innovative because it documents a first attempt to explore patterns of school GCSE attainment at the subject area level in order to investigate whether there are distinct groups of pupils with 'middle' levels of attainment. In previous analysis Connelly, Murray, and Gayle (2013) and Gayle, Murray, and Connelly (2014) did not detect any clear boundaries that demark 'middle' level of attainment in overall or agglomerate measures of school GCSE attainment. We agree with their conclusion that agglomerate school GCSE attainment is best understood as being located on a continuum. We also agree with their conclusion that for many analyses the use of categorical agglomerate measures of GCSE attainment will be suitable, but in such analyses the measures should be considered as coarse groupings of a finer continuous scale rather than substantively distinctive categories.

In the analyses presented above, we have turned the analytical attention away from agglomerate measures of attainment and focused on patterns of attainment at the subject area level. Through a latent variable approach, we identified four distinctive latent educational groups. Two of these latent educational groups were characterised by 'middle' levels of overall (or agglomerate) school GCSE attainment. Therefore, we conclude that this provides some empirical proof of the theoretical conception of 'middle' levels of school GCSE attainment at the subject area level.

Parental socioeconomic position is the most important determinant of overall or agglomerate school GCSE attainment and of latent educational group membership. Pupils from more advantaged socioeconomic backgrounds are more likely to be in the 'Good Grades' group. It is notable that the 'Poor Grades' group comprise a high proportion of

pupils that are from the manual and routine socioeconomic groups. An important finding is that there are no clear socioeconomic differences between pupils in the 'Science' and the 'Arts' latent groups.

Raising levels of school attainment has been part of an ongoing agenda for successive British governments. The identification of four latent educational groups is important. Whilst there is always room for improvement, the 'Good Grades' group perform well across a range of agglomerate measures. They also achieve A–C grades in English and Maths which are increasingly being regarded as priority areas. In stark contrast, a dramatic improvement in school GCSE attainment would be required to move the 'Poor Grades' group up to the benchmark of 5+ GCSEs at grades A–C. This group also perform especially poorly in both English and Mathematics.

The identification of two latent educational groups with 'middle' levels of school GCSE attainment is especially thought provoking. From the evidence presented in this paper, it is clear that in order for pupils in the 'middle' groups to reach the benchmark of 5+ GCSEs with higher grades, there would have to be substantial improvements in some aspects of their performance. Most notably pupils in the 'Science' group require improvements in English, and many would also require improvements in Maths. By contrast, pupils in the 'Arts' group would require less improvement in English. These pupils require greater improvement in Maths however, and their performance in science also requires improvement.

We have reflected on the methodological issue of how to measure school GCSE attainment in the absence of an agreed standard. Pupils study for multiple GCSEs, which include core subjects and non-core subjects, and which are drawn from a wide menu of choices. There is a large array of possible GCSE subject combinations. We have shown that results in individual GCSE subjects are highly correlated. Taken together these two points appeal to the adoption of a latent variable approach because it handles the messy nature of the data whilst not trivialising its complexity. We have demonstrated that a latent variable approach is practicable with large-scale social survey data.

The organisation of the most frequently studied GCSEs into five obvious areas of the school curriculum was a pragmatic attempt to operationalise the latent variable models. We are conscious that other subject area groupings could have been chosen. We are also aware that recording GCSEs in terms of higher and lower grades is one of many possible ways to characterise levels of attainment. We acknowledge that the results partially rest on the way that GCSE results have been grouped and categorised, but we are confident that we have chosen a sensible strategy given the extended exploration previously undertaken as part of the empirical work reported in Playford (2011). We caution that all latent variables approaches to analysing GCSE attainment should consider the effects of grouping and classifying given the absence a universally accepted method of measuring attainment.

In order to test the proof of concept, we restricted our analyses to a single YCS cohort. YCS Cohort 6 was a sensible starting point because GCSEs had been undertaken for a number of years and initial problems associated with teaching and examinations should have been overcome. We are aware that there have been continual improvements in school GCSE attainment (see Department for Education and Skills 2007; House of Commons Education Committee 2014; and Department for Education 2015). Therefore, we are keen to undertake similar analyses with data from more recent school year cohorts in order to establish whether or not 'empirical regularities' exist (see Goldthorpe 2000).

Roberts and MacDonald (2013) remind us of Phil Brown's pithy statement that there is an invisible majority of ordinary pupils who neither leave their names engraved on the school honours board nor gouged into the top of their desks. We conjecture that such

pupils are most likely to be found within the two 'middle' latent educational groups. We see no obvious reasons why school exam results will not continue to be an annual newsworthy item. We foresee that the media focus is most likely to remain on pupils with exceptional levels of performance rather than those with the more modest results that characterise the two 'middle' latent educational groups.

Noah and Eckstein (1992) highlight that over the post-war period there have been various changes to school qualifications in England and Wales; however, the underlying social and educational significance of school qualifications has been preserved. In England, the school leaving age is in the process of being raised (Meyer 2011). School qualifications are likely to undergo changes in the near future, but from current statements we do not envisage that these changes will be far-reaching enough to radically alter the social significance of school qualifications. Indeed, under current plans pupils will still undertake a wide diet of subjects in Year 11.[14]

At the current time a new GCSE grading scheme is likely to be introduced from August 2017. A new set of grades ranging from 1 to 9 (which 9 being the highest) will replace the A*–G scheme.[15] Early indications suggest that the older eight alphabetical grades (A*–G) will not map directly onto the new 1–9 grades, but there will be some general equivalence. Despite the potential reorganisation of GCSEs, and the proposed changes in the grading system, there is no obvious reason to suspect that attainment across GCSE subjects will not continue to be correlated. Therefore, approaches that handle the messy nature of results data whilst not trivialising the underlying complexity of the data will be equally appealing for the analysis of more contemporaneous educational cohorts.

Notes

1. See http://www.education.gov.uk/childrenandyoungpeople/youngpeople/qandlearning/a0074953/review-of-vocational-education-the-wolf-report.
2. Volume 18(1) 2013.
3. Hansen and Vignoles (2005), UCAS (2014) and Kogan (2015) provide an extended account of British qualifications in an international context which some readers may find informative.
4. See also http://www.educationuk.org/global/articles/16-and-under-subjects-qualifications/
5. More recently, School league tables have included the English Baccalaureate (EBacc) which is a performance indicator linked to GCSEs and measures the percentage of pupils in a school who achieve grades A*–C in English, Mathematics, sciences, a foreign language and either history or geography (see https://www.gov.uk/english-baccalaureate-information-for-schools).
6. The A* grade was introduced after the pupils in YCS 6 completed their Year 11 GCSEs. When constructing a measure of GCSE attainment that spanned the introduction of the A* grade Yang and Woodhouse (2001) adopt the strategy of awarding 7 points to both grades A* and A.
7. See http://www.education.gov.uk/schools/performance/secondary_11/PointsScoreAllocation2011.pdf.
8. We also note that Meyer (2011) provides evidence that a large proportion of pupils achieve results spanning at most two grades regardless of the number of GCSEs that they undertake. This further persuades us that a binary measure is appropriate for the analysis.
9. For a full discussion of this measure see Edwards and Edwards (1984).
10. The models were estimated using LEM (Vermunt 1997).
11. See Payne, Payne, and Heath (1994).
12. The results of this model are available on request.

13. 1–4 GCSEs at A–C grades $p = .10$; No GCSE passes $p = .73$. The full results of this model are available on request.
14. See Written Statement to Parliament 9th April 2014 Education Secretary Michael Gove's statement about the publication of reformed GCSE and A' level content (https://www.gov.uk/government/speeches/gcse-and-a-level-reform?utm_source=rss&utm_medium=rss&utm_campaign=statement-to-parliament-gcse-and-a-level-reform) and Department for Education (2010).
15. See http://www.aqa.org.uk/supporting-education/policy/gcse-and-a-level-changes/structure-of-new-gcsesandexamchange.org.uk.

References

Adey, K., and M. Biddulph. 2001. "The Influence of Pupil Perceptions on Subject Choice at 14+ in Geography and History." *Educational Studies* 27 (4): 439–450. doi:10.1080/03055690120071894.

Ashford, S., J. Gray, and M. Tranmer. 1993. *The Introduction of GCSE Exams and Changes in Post-16 Participation*. Employment Department Research Series YCS Reports 23. Sheffield: Employment Department.

Babb, P. 2005. *A Summary of Focus on Social Inequalities*. London: Office for National Statistics.

Bartholomew, D. J., F. Steele, I. Moustaki, and J. I. Galbraith. 2008. *Analysis of Multivariate Social Science Data*. Boca Raton, FL: CRC Press.

Becker, M. P., and I. Yang. 1998. "Latent Class Marginal Models for Cross-Classifications of Counts." *Sociological Methodology* 28 (1): 293–325. doi:10.1111/0081–1750.00050.

Brown, P. 1987. *Schooling Ordinary Kids; Inequality, Unemployment and the New Vocationalism*. London: Tavistock.

Chalabi, M. 2013. "GCSE results 2013: the complete breakdown." *The Guardian*, 22 August. Accessed July 30, 2014. http://www.theguardian.com/news/datablog/2013/aug/22/gcse-results-2013-the-complete-breakdown.

Clark, A. 1995. "Boys into Modern Languages: An Investigation of the Discrepancy in Attitudes and Performance between Boys and Girls in Modern Languages." *Gender and Education* 7 (3): 315–326. doi:10.1080/09540259550039022.

Coles, B. 1986. "Gonna Tear Your Playhouse Down: Towards Reconstructing a Sociology of Youth." *Social Science Teacher* 15: 78–80.

Connolly, P. 2006. "The Effects of Social Class and Ethnicity on Gender Differences in GSCE Attainment: A Secondary Analysis of the Youth Cohort Study of England and Wales 1997–2001." *British Educational Research Journal* 32: 3–21. doi:10.1080/01411920500401963

Connelly, R., S. Murray, and V. Gayle. 2013. "Young People and School GCSE Attainment: Exploring the 'Middle'." *Sociological Research Online* 18 (1): 6. doi:10.5153/sro.2880.

Croxford, L. 2006. *The Youth Cohort Surveys - How Good is the Evidence? Special CES Briefing No. 38*. Edinburgh: Centre for Educational Sociology.

Croxford, L., C. Ianelli, and M. Shapira. 2007. "Documentation of the Youth Cohort Time-Series Datasets." In *UK Data Archive Study Number 5765 - Youth Cohort Time Series for England, Wales and Scotland*, 1984–2002. Edinburgh: UK Data Archive, Centre for Educational Sociology. http://www.esds.ac.uk/doc/5765/mrdoc/pdf/5765userguide.pdf.

Demack, S., D. Drew, and M. Grimsley. 2000. "Minding the Gap: Ethnic, Gender and Social Class Differences in Attainment at 16, 1988–95." *Race, Ethnicity and Education* 3 (2): 117–143. doi:10.1080/13613320050074005

Department for Education. 1985. *General Certificate of Secondary Education: A General Introduction*. London: Her Majesty's Stationery Office.

Department for Education. 2010. *The Importance of Teaching: The Schools White Paper 2010*. London: Department for Education.

Department for Education. 2015. *GCSE and Equivalent Attainment by Pupil Characteristics, 2013 to 2014* (revised). London: Department for Education.

Department for Education and Skills. 2007. *Gender and Education: The Evidence on Pupils in England*. Nottingham: Department for Education and Skills.

Drew, D. 1995. '*Race*', *Education and Work: The Statistics of Inequality*. Aldershot: Avebury.

Drew, D., J. Gray, and N. Sime. 1992. "Against the Odds: The Education and Labour Market Experiences of Black Young People." *Youth Cohort Series 19 Employment Department Training Research and Development Series* 68: 1–56.

Edwards, J. H., and A. W. F. Edwards. 1984. "Approximating the Tetrachoric Correlation Coefficient." *Biometrics* 40 (2): 563.

France, A. 2007. *Understanding Youth in Late Modernity*. Buckingham: Open University Press.

Garner, R. 2013. "Pupils Who'fail' GCSE English or Maths Will be Forced to Stay on at School and Resit their Exams." *The Independent*. 2 September. Accessed July 30, 2014. http://www.independent.co.uk/news/education/education-news/pupils-who-fail-gcse-english-or-maths-will-be-forced-to-stay-on-at-school-and-resit-their-exams-8793614.html.

Gayle, V., D. Berridge, and R. B. Davies. 2003. *Econometric Analysis of the Demand for Higher Education*. Nottingham: DfES Publications.

Gayle, V., and P. S. Lambert. 2007. "Using Quasi-variance to Communicate Sociological Results from Statistical Models." *Sociology* 41 (6): 1191–1208. doi:10.1177/0038038507084830.

Gayle, V., P. S. Lambert, and S. Murray. 2009a. "Growing Up in the 1990s – An exploration of the educational experiences of cohorts of Rising 16s in the BHPS." British Household Panel Survey Conference, University of Essex. July 9–11.

Gayle, V., P. S. Lambert, and S. Murray. 2009b. "School-To-Work in the 1990s: Modelling Transitions with Large-scale Datasets." In *Transitions from Education to Work: New Perspectives from Europe and Beyond*, edited by R. Brooks. Basingstoke: Palgrave Macmillan.

Gayle, V., S. Murray, and R. Connelly. 2014. "Young People and School General Certificate of Secondary Education Attainment: Looking for the 'Missing Middle'." *British Journal of Sociology of Education*. doi:10.1080/01425692.2014.935292.

Gill, T. 2012. *Uptake of GCSE Subjects 2010*. Research Division Assessment Statistics Report Series No.35 – Revised. Cambridge: Cambridge Assessment.

Gillborn, D., and H. S. Mirza. 2000. *Educational Inequality: Mapping Race, Class and Gender*. London: Office for Standards in Education.

Goldthorpe, J. H. 2000. *On Sociology: Numbers, Narratives, and the Integration of Research and Theory*. Oxford: Oxford University Press.

Gorard, S., and C. Taylor. 2002. "What is Segregation? A Comparison of Measures in Terms of 'Strong' and 'Weak' Compositional Invariance." *Sociology* 36 (4): 875–895. doi:10.1177/003803850203600405.

Hansen, K., and A. Vignoles. 2005. "The United Kingdom Education System in a Comparative Context." In *What's the Good of Education? The Economics of Education in the UK*, edited by S. Machin and A. Vignoles, 13–35. Princeton: Princeton University Press.

Haque, Z., and J. F. Bell. 2001. "Evaluating the Performances of Minority Ethnic Pupils in Secondary Schools." *Oxford Review of Education* 27 (3): 357–368. doi:10.1080/03054980120067401.

House of Commons Education Committee. 2014. *Underachievement in Education by White Working Class Children*. London: The Stationery Office Limited.

Institute of Physics. 2012. *It's Different for Girls: The Influence of Schools*. London: Institute of Physics.

Institute of Physics. 2013. *Closing Doors: Exploring Gender and Subject choice in Schools*. London: Institute of Physics.

Jenkins, R. 1983. *Lads, Citizens, and Ordinary Kids: Working Class Youth Lifestyles in Belfast*. London: Routledge and Kegan Paul.

Jin, W., A. Muriel, and L. Sibieta. 2011. *Subject and Course Choices at Ages 14 and 16 amongst Young People in England: Insights from Behavioural Economics.* Research Report DFE-RR160. London: Department for Education.

Joint Council for Qualifications. 2012. *Provisional GCSE Full Course. Results - June 2012 All UK Candidates*. London: Joint Council for Qualifications.

Jones, J., M. Joyce, and J. Thomas. 2003. "Non-employment and labour availability." *Bank of England Quarterly Bulletin*. Autumn: 291–303. http://www.bankofengland.co.uk/publications/Pages/quarterlybulletin/employment.aspx.

Kogan, P. 2015. *British Qualifications 2015: A Complete Guide to Professional, Vocational & Academic Qualifications in the United Kingdom.* 45th ed. London: Kogan Page.

Lamont, A., and K. Maton. 2008. "Choosing Music: Exploratory Studies into the Low Uptake of Music GCSE." *British Journal of Music Education* 25 (3): 267–282. doi:10.1017/S0265051708008103

Leckie, G., and H. Goldstein. 2009. "The Limitations of using School League Tables to Inform School Choice." *Journal of the Royal Statistical Society, Series A Statistics in Society* 172 (4): 835–851. doi:10.1111/j.1467-985X.2009.00597.x

Lim, W. 2013. *A-Level Results 2013 by Subject and Gender*. Accessed July 30. http://blog.concentra.co.uk/2013/08/21/gce-a-level-results-2013/.

Marsh, C., and J. Elliott. 2008. *Exploring Data: An Introduction to Data Analysis for Social Scientists*. Cambridge: Polity Press.

McCutcheon, A. L. 1987. *Latent Class Analysis*. London: Sage.

McCutcheon, A. L. 1996. "Models for the Analysis of Categorically-Scored Panel Data." In *Analysis of Change: Advanced Techniques in Panel Data Analysis*, edited by U. Engel and J. Reinecke, 15–34. Berlin: de Gruyter.

McCutcheon, A. L. 2002. "Basic Concepts and Procedures in Single- and Multiple-Group Latent Class Analysis." In *Applied Latent Class Analysis*, edited by J. A. Hagenaars and A. L. McCutcheon, 56–85. Cambridge: Cambridge University Press.

Merton, R. K. 1987. "Three Fragments from a Sociologist's Notebooks: Establishing the Phenomenon, Specified Ignorance, and Strategic Research Materials." *Annual Review of Sociology* 13 (1): 1–29. doi:10.1146/annurev.so.13.080187.000245.

Meyer, L. 2011. *The Value of GCSEs*. Guildford: Centre for Education Research and Policy.

Mobley, M., C. Emerson, I. Goddard, S. Goodwin, and R. Letch. 1986. *All about GCSE*. London: Heinemann Educational.

Mujtaba, T., and M. J. Reiss. 2013. "What Sort of Girl Wants to Study Physics after the Age of 16?" *International Journal of Science Education* 35 (17): 2979–2998. doi:10.1080/09500693.2012.681076

Murphy, P., and E. Whitelegg. 2006. *Girls in the Physics Classroom: A Review of the Research on the Participation of Girls in Physics*. London: Institute of Physics.

Murray, S. J. 2011. "Growing up in the 1990s: Tracks and Trajectories of the 'Rising 16's': A Longitudinal Analysis using the British Household Panel Survey." PhD thesis. Stirling: University of Stirling.

Noah, H. J., and M. A. Eckstein. 1992. The Two Faces of Examinations: A Comparative and International Perspective. In *Examinations: Comparative and International Studies*, edited by H. J. Noah and M. A. Eckstein. Oxford: Pergamon Press.

North, J. 1987. *The GCSE: An Examination*. London: Claridge Press, in association with Education Research Trust.

Owen-Jackson, G. ed. 2013. *Debates in Design and Technology Education*. Abingdon: Routledge.

Pau, R., W. Hall, and M. Grace. 2011. "'It's Boring': Female Students' Experience of Studying ICT and Computing." *School Science Review* 92: 89–94.

Payne, J. 1995. *Routes beyond Compulsory Schooling.* Employment Department Research Series YCS Reports, 31. Sheffield: Employment Department.

Payne, J. 2000. *Young People not in Education, Employment or Training: Data from the England and Wales Youth Cohort Study*. Research Report 201. Nottingham: Department for Education and Employment.

Payne, J. 2001. *Post-16 Students and Part-Time Jobs: Patterns and Effects.* Research Report 323. London: Department for Education and Skills.

Payne, J. 2003. *Vocational Pathways at Age 16–19*. Research Report 501. Sheffield: Department for Education and Skills.

Payne, C., J. Payne, and A. Heath. 1994. "Modelling Trends in Multiway Tables." In *Analyzing Social and Political Change*, edited by A. Dale and R. B. Davies, 43–74. London: Sage.

Phillips, C. 2009. "Ethnic Inequalities: Another 10 Years of the Same?" In *Towards a More Equal Society? Poverty, Inequality and Policy Since 1997*, edited by J. Hills, T. Sefton, and K. Stewart, 179–200. Bristol: The Policy Press.

Playford, C. J. 2011. "The Role of Families in the Stratification of Attainment: Parental Occupations, Parental Education and Family Structure in the 1990s." PhD thesis. Stirling: University of Stirling.

Prandy, K., M. Unt, and P. S. Lambert. 2004. "Not by degrees: Education and social reproduction in twentieth-century Britain." *International Sociological Association Research Committee 28 on Social Stratification and Mobility*, Neuchatel, Switzerland.

Pye, J. 1988. *Invisible Children: Who are the Real Losers in School?* Oxford: Oxford University Press.

Rice, P. 1999. "The Impact of Local Labour Markets on Investment in Further Education: Evidence from the England and Wales Youth cohort Studies." *Journal of Population Economics* 12 (2): 287–312. doi:10.1007/s001480050100

Roberts, S. 2011. "Beyond 'NEET' and 'Tidy' Pathways: Considering the 'Missing Middle' of Youth Transition Studies." *Journal of Youth Studies* 14 (1): 21–39. doi:10.1080/13676261.2010.489604.

Roberts, S. 2012. "'I Just Got on with it': The Educational Experiences of Ordinary, Yet Overlooked, Boys." *British Journal of Sociology of Education* 33 (2): 203–221. doi:10.1080/01425692.2011.649832

Roberts, S., and R. Macdonald. 2013. "Introduction for Special Section of Sociological Research Online: The Marginalised Mainstream: Making Sense of the'Missing Middle' of Youth Studies." *Sociological Research Online* 18 (1): 21.

Rollock N., D. Gillborn, C. Vincent, and S. Ball. 2015. *The Colour of Class*. London: Routledge.

Schneider, S. 2011. *Measuring Educational Attainment. Survey Question Bank: Topic Overview 6.* Accessed August 18, 2013. http://surveynet.ac.uk/sqb/topics/education/sqb_education_schneider.pdf.

Sullivan, A., A. Heath, and C. Rothon. 2011. "Equalisation or Inflation? Social Class and Gender Differentials in England and Wales." *Oxford Review of Education* 37: 215–240.

Sullivan, A., A. Zimdars, and A. Heath. 2010. "The Social Structure of the 14–16 Curriculum in England." *International Studies in Sociology of Education* 20 (1): 5–21. doi:10.1080/09620211003655572

Taylor, J. 2011. *The English Baccalaureate: How not to Measure School Performance*. Lancaster: Lancaster University Management School.

The Russell Group. 2013. *Informed Choices: A Russell Group Guide to Making Decisions About Post-16 Education*. London: Russell International Excellence Group.

Tukey, J. W. 1977. *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.

Tunstall, R. 2011. *Social Housing and Social Exclusion 2000–2011*. Case Paper 153. London: CASE, LSE.

UCAS. 2014. *UK Qualifications: For Entry to University or College in 2015*. Cheltenham: UCAS.

Vermunt, J. K. 1997. *LEM: A General Program for the Analysis of Categorical Data*. Tilburg: Department of Methodology and Statistics, Tilburg University.

Warmington, P., and P. Murphy. 2004. "Could Do Better? Media Depictions of UK Educational Assessment Results." *Journal of Education Policy* 19 (3): 285–299. doi:10.1080/0268093042000207629.

Webber, R., and T. Butler. 2007. "Classifying Pupils by where they Live: How Well Does this Predict Variations in their GCSE Results?" *Urban Studies* 44 (7): 1229–1254. doi:10.1080/00420980701302353.

Weedon, P. 2011. "Inequality of Opportunity and the Take-up of GCSE Geography." *Curriculum Journal* 22 (3): 401–422. doi:10.1080/09585176.2011.601673.

Wolf, A. 2011. *Review of Vocational Education: The Wolf Report*. London: Department for Education.

Yang, M., and G. Woodhouse. 2001. "Progress from GCSE to A and AS Level: Institutional and Gender Differences, and Trends over Time." *British Educational Research Journal* 27: 245–267.

Data citation

Youth Cohort Study of England and Wales, 1992–1994; Cohort Six, Sweep One to Three [computer file]. Colchester, Essex: UK Data Archive [distributor], August, 1996. SN: 3532, http://dx.doi.org/10.5255/UKDA-SN-3532-1